REMARKS

The present application was filed on October 11, 2000 with claims 1-30. Claims 1, 10, 11, 20, 21 and 30 are the independent claims and have been amended. Claims 1-30 remain pending.

In the outstanding Office Action dated November 4, 2004, the Examiner: (i) objected to the drawings; (ii) rejected claims 1-30 under 35 U.S.C. §101; (iii) rejected claims 1-30 under 35 U.S.C. §112, first paragraph; and (iv) rejected claims 1-3, 10-13, 21-23 and 30 under 35 U.S.C. §102(b) as being anticipated by U.S. Patent No. 5,832,182 to Zhang et al. (hereinafter "Zhang").

With regard to the objection to the drawings, formal drawings were previously filed on December 30, 2003. Copies of the drawings and associated transmittal are again submitted herewith.

With regard to the rejections of claims 1-30 under 35 U.S.C. §101 and §112, first paragraph, Applicants have amended independent claims 1 and 10 so that they are claimed for practice on a computer. Independent claims 11, 20, 21 and 30 recite the use of a processor and/or a machine readable medium having one or more programs, and thereby are already claimed for practice on a computer.

Examiner contends that Applicants' "data set" is an abstract idea. Applicants believe that the Office Action fails to appreciate the nature of the subject matter of the present invention. For example, on page 1, lines 4-7, the application states that the present invention is related to methods for performing "detection in accordance with various high dimensional domain applications where it is important to be able to find and detect outliers which deviate considerably from the rest of the data."

As explained in the Summary of the Invention, at pages 4 and 5 of the present specification, the present invention provides methods and apparatus for outlier detection which find outliers by observing the density distributions of projections from the data. The invention defines outliers for data by looking at projections of the data which have abnormally low density. In general, claims 1-30 are directed toward methods, apparatus, and articles of manufacture for detecting one or more outliers in a data set. In one embodiment, one or more sets of dimensions and corresponding ranges in the data set which are sparse in density are determined, and one or more data points in the data

set which contain these sets of dimensions and corresponding ranges are identified as the one or more outliers in the data set.

As later explained in the Detailed Description at page 12, lines 19-26, FIG. 2 illustrates a hardware implementation suitable for employing outlier detection methodologies according to an embodiment of the present invention.

> It is assumed that the server 30 contains a large repository of data which is used for the purpose of data mining. The requests for finding the outliers along with the corresponding data sets are specified at the client end 10. These requests are then responded to using the methodologies of the present invention as implemented on the server end 30. The computation is performed by the CPU 32. The data on which the analysis is carried out may already be available at the server on its disk 36, or it may be specified by the client. In either case, the computation is performed at the server end, and the results are returned to and presented to (e.g., displayed) the client.

Still further, at page 13, lines 10-27, FIG. 3 illustrates a flow diagram of the overall process for outlier detection according to an embodiment of the present invention.

> In step 310, an encoding is determined for the database by creating the intervals for each dimension. For example, let us consider a 2-dimensional database in which there are two attributes, age and salary. Then, the encoding will be created by a string of length 2. Let us say that each of the two attributes is divided into $p = 3$ ranges. For example:
>
> Age: Range 1→ 0-30
> Range 2→ 31-60
> Range 3→ 61 and above
> Salary: Range 1→ 0-50,000
> Range 2 →50,001-100,000
> Range 3 →100,001 and above
>
> Then, a 28-year-old person with a salary of 60,000 would been encoded as 12 (i.e., Age Range 1 followed by Salary Range 2), whereas a 62-year-old person with a salary of 20,000 would be encoded as 31. Also, in step 310, we first divide each of the attributes of the database into intervals. Each of these intervals is chosen in such a way that an equal number of records satisfy them. Thus, if p intervals are chosen, then exactly a fraction $1/p$ of the records in the database lie in each of these

intervals. In step 320, each record in the database is expressed in terms of these intervals.

It is well-established law that an invention directed toward techniques that may be implemented in one or more processing devices, such as one or more computer systems, as illustrated in FIGS. 2 and 3, is proper statutory subject matter under §101 if it produces a useful, concrete and tangible result. *State Street Bank & Trust Co. v. Signature Financial Group, Inc.*, 47 USPQ2d 1596 (Fed. Cir. 1998). In fact, the Federal Circuit in *State Street* goes on to state that subject matter is rendered statutory even if the useful result is expressed in numbers, such as price, profit, percentage, cost, or loss. Thus, the issue is whether the present invention produces a useful, concrete and tangible result. Applicants strongly assert that the invention does produce a useful, concrete and tangible result. At a minimum, the invention, as recited in the subject claims, produces a useful, concrete and tangible result in the form of one or more data points identified as one or more outliers in the data set.

The invention may be implemented in accordance with computer-based methods (e.g., independent claims 1 and 10), processor-based apparatus (e.g., independent claims 11 and 20), and articles of manufacture comprising a machine readable medium containing one or more programs which when executed implement the various inventive steps (e.g., independent claims 21 and 30). Thus, since the invention is described as being implemented in one of these forms and one or more data points in a data set are identified as one or more outliers of the data set, as described above, Applicants believe that claims 1-30 contain statutory subject matter under §101 and therefore request withdrawal of the rejections under§101 and §112, first paragraph.

With regard to the rejection of claims 1-3, 10-13, 21-23 and 30 under 35 U.S.C. §102(b) as being anticipated by Zhang, Applicants have amended independent claims 1, 10, 11, 20, 21 and 30 to recite that the determination of subsets of dimensions and corresponding ranges, or patterns, in the data set which are sparse in density use an algorithm capable of utilizing at least one of the processes of solution recombination, selection and mutation over a population of multiple solutions. Support for the amendment may be found throughout the present specification, e.g., see page 5, lines 12-18. Zhang discloses methods for clustering very large databases but fails to disclose such an

algorithm, as described in independent claims 1, 10, 11, 20, 21 and 30, for the determination of subsets of dimensions and corresponding ranges in the data set. Dependent claims 2, 3, 12, 13, 22 and 23 are patentable at least by virtue of their dependency from independent claims 1, 11 and 21. Dependent claims 2, 3, 12, 13, 22 and 23 also recite patentable subject matter in their own right. Accordingly, withdrawal of the rejection to claims 1-3, 10-13, 21-23 and 30 under 35 U.S.C. §102(b) is therefore respectfully requested.

In view of the above, Applicants believe that claims 1-30 are in condition for allowance, and respectfully request withdrawal of the §101, §112 and §102(b) rejections.

Respectfully submitted,

Date: February 4, 2005

Robert W. Griffith
Attorney for Applicant(s)
Reg. No. 48,956
Ryan, Mason & Lewis, LLP
90 Forest Avenue
Locust Valley, NY 11560
(516) 759-4547